

ANALISIS BUTIR SOAL SEBAGAI STANDARISASI MUTU SEKOLAH DASAR PADA MATA PELAJARAN PAI DENGAN MENGGUNAKAN PENDEKATAN RASCH MODEL DI SD MA'ARIF NU KECAMATAN PANDAAN PASURUAN

Askhabul Kirom, Mochamad Hasyim

Universitas Yudharta Pasuruan, East Java, Indonesia

k1r0m@yudharta.ac.id, hasyim@yudharta.ac.id

ABSTRACT:

Received:
April 14th 2021

Revised:
May 25th 2021

Accepted:
June 11th 2021

Comprehensive tests aim to improve the quality of assessment to determine the item's level of difficulty, know the level of conformity of the item (item fit), and know the existence of biased question items. Explorative, descriptive research is a post-hoc diagnosis described as a retro-fitting approach. The strengthening method is carried out by analyzing question items and grain response data to find quality information about students in the final semester 1 of Islamic Religious Education subjects in the 2020/2021 school year. The study's findings are as follows: 1. The difficulty level of the item measure seen in S14 question item of the Measure value is +2.82, and only 25 students answer correctly. 2. The level of conformity of the item question (item fit) there is one response that is outside the boundary of the Outfit trust room that is in the item question S14, other than the item of the question has met the criteria of outfit α -standard, outfit means-square value, and point measure correlation. 3. Detection of biased question items found in question item S13, where the item question S13 probability value is 0.02, which is below 5% (0.05).

Keywords:

Analisis Butir; Tes; Rasch Model

PENDAHULUAN

Tes adalah Prosedur yang harus diperhatikan dalam rangka pengukuran atau cara yang dapat dipergunakan serta assement di bidang pendidikan. Keterampilan yang harus dikuasai oleh seorang pendidik adalah tentang sistem penilaian hasil belajar Peserta Didik. Dalam hasil belajar siswa dan penilaian proses di lembaga pendidikan, aspek yang berkaitan dengan pemilihan alat assement, diantaranya; analisis butir soal taupun penyusunan soal, hal ini dilakukan guna memperoleh kualitas soal yang memadai, serta interpretasi data dari hasil penilaian.¹

Tes bertujuan untuk proses skrining awal, remidi, atau tes untuk mengetahui manakah butir soal yang tergolong sulit, sangat sulit dan mana yang mudah atau sangat mudah untuk dikerjakan. Tes dilakukan untuk melihat kesesuaian butir (item fit) atau kualitas kesesuaian butir dengan model, , serta mendeteksi adanya butir soal yang bias. Hasil tes memberikan informasi yang bisa digunakan untuk melihat suatu bentuk tes standar atau belum, tes standar ini perlu dirancang dengan baik untuk dikenal

¹ Abdul Kadir, "Menyusun Dan Menganalisis Tes Hasil Belajar," *AL-TA'DIB: Jurnal Kajian Ilmu Kependidikan*, 8, No. 2 (2015): 70–81.

secara luas. Bentuk tes standar yang dirancang dalam penelitian ini adalah item soal pilihan ganda dengan satuan pilihan jawaban benar serta siswa disediakan empat pilihan jawaban.

Kemampuan seorang guru dalam membuat tes dengan tujuan untuk meningkatkan kompetensi siswa sangat dibutuhkan, dengan demikian, tes blok, apakah tes pada setiap kompetensi dasar ataupun tes ujian akhir yang dilakukan oleh para guru haruslah tes yang dikategorikan standard. Tes yang berada di sekolah masih berupa *Item Pool* (tes tersebut belum diketahui karakteristiknya dan tes yang belum dikatakan standard, tingkat kesulitan butir, baik daya beda butir maupun tebaan butir) belum berupa *Item Bank* (tes tersebut sudah diketahui karakteristiknya dan tes yang standard, baik daya beda soal, tebakan, maupun tingkat kesulitan). Tes yang belum standar akan memberikan informasi berkaitan dengan kemampuan siswa yang bias dan juga tidak akurat, sehingga data maupun informasi yang akan didapat masih diragukan akan kebenarannya.

Untuk mengetahui tingkat kesulitan butir soal (*Item Measure*), mengetahui tingkat kesesuaian butir soal (*Item Fit*), serta mengetahui adanya butir soal yang bias. Penulis menggunakan pendekatan rasch model di Sekolah Dasar Ma'arif NU di Kecamatan Pandaan Kabupaten Pasuruan. Hal ini dilakukan karena sejauh ini para guru mata pelajaran belum melakukan tes yang bersifat komprehensif untuk peningkatan kualitas penilaian, sehingga belum diketahui kualitas soal, tingkat kesulitan soal sampai kepada deteksi miskonsepsi, adanya bias soal. Senada dengan hasil penelitian Isnaini, dkk. bahwa soal tes dengan menggunakan Rasch model dapat menunjukkan tes yang baik serta mengetahui kemampuan siswa dapat diperkirakan dengan tepat.²

Berdasarkan uraian tersebut maka penulis menganggap perlunya diadakan penelitian tentang analisis butir soal sebagai tes standarisasi mutu Sekolah Dasar pada mata pelajaran PAI dengan menggunakan pendekatan Rasch model di SD Ma'arif NU Kecamatan Pandaan Kabupaten Pasuruan.

METODE

Penelitian ini merupakan penelitian deskriptif eksploratif terhadap kumpulan data item dan tanggapan siswa SD Ma'arif NU di Kecamatan Pandaan Kabupaten Pasuruan semester I mata pelajaran Pendidikan Agama Islam Tahun Pelajaran 2020/2021. Penelitian ini menggunakan *post-hoc diagnostic* yang diilustrasikan sebagai pendekatan *retro-fitting*, pendekatan dilakukan melalui analisis butir soal untuk menemukan informasi kualitas soal siswa kelas V pada ujian akhir semester I mata pelajaran pendidikan agama Islam pada tahun pelajaran 2020/2021.

Subjek dalam penelitian ini adalah Siswa kelas V yang telah mengerjakan soal tes standarisasi mutu sekolah dasar semester I mata pelajaran PAI yang diujikan di SD Ma'arif Sumberejo, SD Ma'arif NU Ngampir, dan SD Ma'arif Riyadul Arkham Kecamatan Pandaan Kabupaten Pasuruan. Adapun

² Isnani Isnani et al., "Estimation of College Students' Ability on Real Analysis Course Using Rasch Model," *RESEARCH AND EVALUATION IN EDUCATION* 5, No. 2 (2019): 95–102.

yang dianalisis dalam hal ini adalah butir soal pilihan ganda sejumlah 20 soal dengan 4 pilihan jawaban yaitu A, B, C, D.

Pengumpulan data dalam penelitian ini berupa dokumen yaitu mengumpulkan data mengenai lembar jawaban siswa kelas V di SD Ma’arif Sumberejo, SD Ma’arif NU Ngampir, dan SD Ma’arif Riyadul Arkham Kecamatan Pandaan Kabupaten Pasuruan.

Adapun dokumentasi yang digunakan untuk pengumpulan data dalam penelitian yaitu:

- 1) Lembar soal siswa terhadap tes standarisasi mutu sekolah dasar semester I mata pelajaran pendidikan agama Islam kelas V (lima) tahun pelajaran 2020-2021.
- 2) Lembar jawaban siswa terhadap soal tes standarisasi mutu sekolah dasar semester I mata pelajaran PAI kelas V (lima) tahun pelajaran 2020-2021.

Analisa data digunakan untuk memproses, menyusun dan mengolah data yang telah sedemikian rupa sehingga dapat dipahami. Adapun analisa data dalam penelitian ini menggunakan *Software* Ministep. Ministep merupakan program computer yang dapat bekerja khusus untuk analisis pemodelan *Rasch* dengan sistem Ms. Windows yang dibuat oleh John linacre.

HASIL DAN PEMBAHASAN

1. Tingkat kesulitan butir soal (*item measure*)

Tingkat kesulitan pada butir soal (*item measure*) atau bisa disebut *item logit* dapat diketahui dengan menggunakan program Ministep yaitu dipilih pada table ke-4 *item measure*. Pada table 1 terlihat adanya beberapa kolom yang memberikan data informasi tentang tingkat kesulitan tentang tiap butir soal (*item measure*), sebagaimana pada tabel sebagaimana di bawah ini:

Tabel 1. Item Measure

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
14	25	81	2.82	.28	1.26	1.9	2.38	3.3	.31	.51	67.9	76.5	S14
10	36	81	2.00	.27	1.08	.7	1.17	.8	.48	.53	74.4	72.4	S10
19	38	81	1.86	.27	1.10	.9	1.28	1.4	.47	.53	67.9	72.4	S19
13	41	81	1.65	.26	1.01	.1	.86	-.7	.55	.54	66.7	72.3	S13
6	46	81	1.30	.27	.93	-.5	.98	.0	.56	.53	78.2	73.1	S6
3	48	81	1.16	.27	1.21	1.7	1.29	1.5	.41	.53	70.5	73.6	S3
12	49	81	1.08	.27	1.02	.2	.99	.0	.52	.53	74.4	73.8	S12
9	55	81	.63	.28	.98	-.1	.84	-.6	.54	.52	74.4	76.6	S9
2	59	81	.30	.29	1.14	1.0	1.01	.1	.44	.50	73.1	79.1	S2
11	60	81	.21	.30	1.02	.2	1.07	.3	.47	.49	82.1	79.7	S11
4	61	81	.12	.30	1.05	.4	1.29	1.0	.45	.49	78.2	80.3	S4
20	66	81	-.38	.33	.80	-1.1	.67	-.8	.55	.45	87.2	83.7	S20
7	67	81	-.49	.34	1.11	.6	1.09	.4	.39	.44	80.8	84.6	S7
16	69	81	-.73	.36	.77	-1.1	.46	-1.2	.56	.42	91.0	86.4	S16
17	72	81	-1.14	.39	.56	-2.0	.24	-1.7	.61	.38	94.9	88.9	S17
18	73	81	-1.31	.41	.80	-.7	.33	-1.2	.50	.37	88.5	89.9	S18
15	75	81	-1.69	.46	.83	-.5	.99	.3	.37	.33	92.3	92.3	S15
5	76	81	-1.91	.50	.85	-.3	.27	-.8	.42	.30	93.6	93.5	S5
1	78	81	-2.52	.62	.93	.0	.69	.0	.27	.24	96.2	96.1	S1
8	79	81	-2.97	.74	.80	-.1	.15	-.8	.33	.20	97.4	97.4	S8
MEAN	58.6	81.0	.00	.36	.96	.1	.90	.1			81.5	82.1	
S.D.	15.4	.0	1.56	.13	.17	.9	.49	1.1			10.1	8.3	

Pada tabel tersebut terlihat tingkat kesulitan butir soal (*item measure*) yaitu dengan melihat *item measure*. Nilai *Measure* yang tinggi hal tersebut menunjukkan butir soal yang tinggi tingkat

kesulitan soalnya. Hal ini juga sejalan dengan *Total Score* yang menyatakan dengan jumlah jawaban benar. Adapun tingkat kesulitan butir soal (*Item measure*) terlihat pada butir soal ke-14 nilai *Measure*-nya adalah +2.82 dan hanya 25 siswa yang menjawab dengan benar (lihat pada kolom *total score*).

Dari tabel data *Measure* ini hal lain yang bisa dijelaskan adalah berkaitan dengan tingkat kesulitan, jika dibandingkan butir soal (S10) dengan butir soal (S9), bisa disimpulkan bahwa pada butir soal (S10) tiga (3) kali tingkat kesulitannya dengan butir soal (S9). Sedikit berbeda jika dibandingkan dari skor yang diperoleh, jawaban yang benar ada 36 siswa untuk butir soal (S10) jumlah, sedangkan ada (55) siswa untuk soal (S9).

Pada *Item Measure* juga dapat memberikan informasi untuk pengelompokan tingkat kesulitan butir soal (*Item Measure*) yaitu melihat nilai *Standar Deviasi* (1.71). Jika nilai ini dikombinasikan dengan rata-rata *Measure*, artinya, ada empat kelompok soal berdasarkan tingkat kesulitannya, dalam hal ini tingkat kesulitan butir soal dikelompokkan menjadi empat kelompok (sangat sulit, sulit, sangat mudah, dan mudah). Untuk kelompok butir soal sulit pada soal (S19, S13, S6, S3, dan S12) dengan ketentuan (0,0 logit+1SD), kelompok butir soal kategori tingkatan sangat sulit pada soal (S14, dan S10) dengan ketentuan (lebih besar dari +1SD), tingkat pengelompokan butir soal mudah pada soal (S9, S2, S11, S4, S20, S7, S16, S17, S18, S15, dan S5) dengan ketentuan (0,0 Logit-1SD), sedangkan pengelompokan tingkat butir soal sangat mudah terdapat pada soal (S1, dan S8) dengan ketentuan (lebih kecil dari -1SD).

2. Tingkat kesesuaian butir soal (*Item Fit*)

Penjelasan selanjutnya adalah melihat *Item Fit* dengan pemodelan rasch. *item fit* menjelaskan apakah butir soal berfungsi untuk mengetahui tingkat kesesuaian pada butir soal secara normal ataupun tidak. Jika didapati butir soal yang tampak tidak *fit*, hal tersebut tampak adanya *miskonsepsi* pada butir soal tersebut. Data tersebut sangat bermanfaat bagi guru untuk membenahi kualitas butir soal sehingga *miskonsepsi* bisa terhindari pada saat merancang butir soal itu kembali. Di bawah ini tabel *item fit order* secara berurutan butir soal (*item measure*):

Tabel 2. Item Fit

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
14	25	81	2.82	.28	1.26	1.9	2.38	3.3	A .31	.51	67.9	76.5	S14
3	48	81	1.16	.27	1.21	1.7	1.29	1.5	B .41	.53	70.5	73.6	S3
4	61	81	.12	.30	1.05	.4	1.29	1.0	C .45	.49	78.2	80.3	S4
19	38	81	1.86	.27	1.10	.9	1.28	1.4	D .47	.53	67.9	72.4	S19
10	36	81	2.00	.27	1.08	.7	1.17	.8	E .48	.53	74.4	72.4	S10
2	59	81	-.30	.29	1.14	1.0	1.01	.1	F .44	.50	73.1	79.1	S2
7	67	81	-.49	.34	1.11	.6	1.09	.4	G .39	.44	80.8	84.6	S7
11	60	81	-.21	.30	1.02	.2	1.07	.3	H .47	.49	82.1	79.7	S11
12	49	81	1.08	.27	1.02	.2	.99	.0	I .52	.53	74.4	73.8	S12
13	41	81	1.65	.26	1.01	-.1	.86	-.7	J .55	.54	66.7	72.3	S13
15	75	81	-1.69	.46	.83	-.5	.99	-.3	J .37	.33	92.3	92.3	S15
9	55	81	.63	.28	.98	-.1	.84	-.6	I .54	.52	74.4	76.6	S9
6	46	81	1.30	.27	.93	-.5	.98	.0	H .56	.53	78.2	73.1	S6
1	78	81	-2.52	.62	.93	.0	.69	.0	G .27	.24	96.2	96.1	S1
5	76	81	-1.91	.50	.85	-.3	.27	-.8	F .42	.30	93.6	93.5	S5
8	79	81	-2.97	.74	.80	-.1	.15	-.8	E .33	.20	97.4	97.4	S8
20	66	81	-.38	.33	.80	-1.1	.67	-.8	D .55	.45	87.2	83.7	S20
18	73	81	-1.31	.41	.80	-.7	.33	-1.2	C .50	.37	88.5	89.9	S18
16	69	81	-.73	.36	.77	-1.1	.46	-1.2	B .56	.42	91.0	86.4	S16
17	72	81	-1.14	.39	.56	-2.0	.24	-1.7	A .61	.38	94.9	88.9	S17
MEAN	58.6	81.0	.00	.36	.96	.1	.90	.1			81.5	82.1	
S.D.	15.4	.0	1.56	.13	.17	.9	.49	1.1			10.1	8.3	

Menurut Boone³ serta Bond dan Fox⁴, bahwa nilai *outfit z-standard*, *outfit means square*, serta nilai *point measure correlation* dikatakan diterima apabila tingkat kesesuaian butir soal (*item fit*) sudah memenuhi kriteria yang digunakan. Apabila pada ke_tiga kriteria tersebut belum terpenuhi pada butir soal, dapat dipastikan butir soalnya tidak diterima sehingga perlu adanya perbaikan. Hal ini dilakukan untuk dapat menjamin di lain waktu adanya kesesuaian butir soal yang sesuai serta berkualitas.

Menurut Boone⁵, untuk memverifikasi kesesuaian butir soal yang tidak diterima (misfit atau outliers) kriteria yang digunakan adalah:

- a. Nilai ZSTD diterima jika $-2,0 < ZSTD < + 2$,
- b. Nilai MNSQ diterima jika $0,5 < MNSQ < 1,5$
- c. Nilai Pt Mean Corr diterima jika $0,4 < Pt Measure Corr < 0,85$

Tabel 2 menunjukkan bahwa butir soal yang kurang bagus pada butir soal S14 sehingga perlu diperbaiki butir soal tersebut, hal ini dikarenakan adanya ketidaksesuaian kriteria nilai ZSTD, MNSQ dan Pt. Mean Corr yang tidak terpenuhi kriterianya. Sehingga kesimpulan akhirnya hanya pada butir soal S14 yang perlu diperbaiki.

3. Deteksi adanya butir soal yang bias

Jika dideteksi siswa dengan karakter tertentu makin diuntungkan dibandingkan siswa lainnya hal itu menunjukkan butir soal dikatakan mengandung bias. Apabila butir soal lebih mudah untuk dijawab oleh siswa perempuan daripada siswa laki-laki, maka hal ini menunjukkan adanya butir soal bisa dikatakan bias gender. Dalam hal untuk mendeteksi bias butir soal ini bisa dilihat dengan deteksi DIF (*Differential Item Functioning*) atau keberfungsian butir diferensial. Seperti halnya pada tabel di bawah ini:

Tabel 3. Deteksi butir soal yang bias

DIF class specification is: DIF=\$S3W1

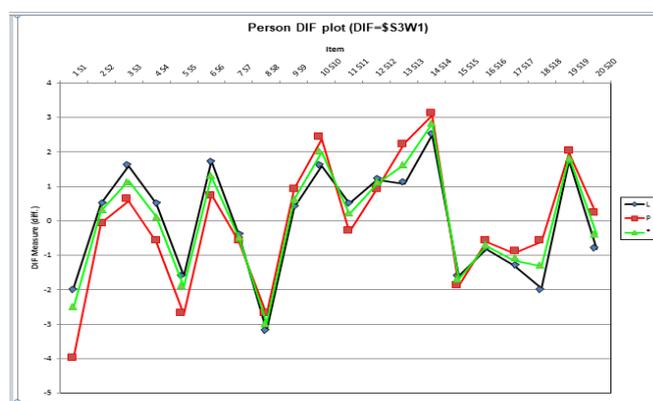
Person CLASS	DIF MEASURE	DIF S.E.	Person CLASS	DIF MEASURE	DIF S.E.	CONTRAST	JOINT S.E.	t	d.f.	Prob.	Mantel-Haenszel Chi-squ	Size Prob.	Item Number	Name	
L	1.99	.64	P	-3.99	1.84	2.04	1.95	1.05	62	.2992	5000	.4795	1	S1	
L	1.52	.38	P	-.06	.49	-.59	.62	-.95	75	.3465	3411	.5592	.58	2	S2
L	1.68	.37	P	-.57	.44	-1.03	.57	-1.80	75	.0753	1.4555	.2276	.77	3	S3
L	1.52	.38	P	-.61	.56	-1.13	.67	-1.68	73	.0963	1.3251	.2497	1.15	4	S4
L	1.59	.37	P	-.71	1.07	-1.12	1.21	-.92	78	.3583	5000	.4795	.80	5	S5
L	1.74	.37	P	-.76	.42	-.98	.56	-1.75	75	.0848	4.6862	.0304	1.59	6	S6
L	-3.42	.43	P	-.61	.56	-.19	.70	-.27	75	.7913	4994	.4798	.78	7	S7
L	1.60	.37	P	-.71	1.07	-.49	1.49	-.33	75	.7442	3333	.5637	1.00	8	S8
L	1.38	.38	P	-.93	.41	-.55	.56	-.99	75	.3278	.8992	.3430	1.80	9	S9
L	1.20	.37	P	2.38	.37	-.78	.52	-1.49	75	.1392	1.4248	.2326	1.10	10	S10
L	1.52	.38	P	-.32	.52	-.84	.64	-1.31	74	.1935	.9775	.3228	.87	11	S11
L	1.07	.37	P	2.25	.37	-1.18	.52	-2.28	75	.0255	5.6259	.0177	2.08	13	S13
L	2.50	.41	P	3.07	.38	-.57	.56	-1.01	75	.3147	.0882	.9279	.12	14	S14
L	1.59	.37	P	-1.88	.79	-.29	.98	-.30	74	.7668	6000	.4795	1.17	15	S15
L	-.81	.46	P	-.61	.56	-.21	.72	-.29	75	.7749	.0744	.7850	1.17	16	S16
L	1.29	.52	P	-1.94	.60	-.35	.80	-.44	75	.6603	.7500	.3865	.22	17	S17
L	-1.95	.64	P	-.61	.56	-1.34	.85	-1.59	75	.1165	3.5854	.0583	1.18	18	S18
L	1.74	.37	P	1.98	.37	-.23	.53	-.45	75	.6570	.0886	.9262	.22	19	S19
L	-.81	.46	P	-.17	.47	-.98	.66	-1.49	75	.3405	.9655	.3258	1.06	20	S20
P	3.99	1.84	L	1.99	.64	2.04	1.95	1.05	62	.2992	5000	.4795	1	S1	
P	-.06	.49	L	1.52	.38	-.59	.62	-.95	75	.3465	3411	.5592	.58	2	S2
P	-.57	.44	L	1.68	.37	-1.03	.57	-1.80	75	.0753	1.4555	.2276	.77	3	S3
P	-.61	.56	L	1.52	.38	-1.13	.67	-1.68	73	.0963	1.3251	.2497	1.15	4	S4
P	-.71	1.07	L	1.74	.37	-1.12	1.21	-.92	78	.3583	5000	.4795	.80	5	S5
P	-.76	.42	L	1.74	.37	-.98	.56	-1.75	75	.0848	4.6862	.0304	1.59	6	S6
P	-3.42	.43	L	-.61	.56	-.19	.70	-.27	75	.7913	4994	.4798	.78	7	S7
P	-.71	1.07	L	-.71	1.07	-.49	1.49	-.33	75	.7442	3333	.5637	1.00	8	S8
P	-.93	.41	L	-.93	.41	-.55	.56	-.99	75	.3278	.8992	.3430	1.80	9	S9
P	2.38	.37	L	2.38	.37	-.78	.52	-1.49	75	.1392	1.4248	.2326	1.10	10	S10
P	-.32	.52	L	1.52	.38	-.84	.64	-1.31	74	.1935	.9775	.3228	.87	11	S11
P	2.25	.37	L	1.07	.37	-1.18	.52	-2.28	75	.0255	5.6259	.0177	2.08	13	S13
P	3.07	.38	L	2.50	.41	-.57	.56	-1.01	75	.3147	.0882	.9279	.12	14	S14
P	-1.88	.79	L	1.59	.37	-.29	.98	-.30	74	.7668	6000	.4795	1.17	15	S15
P	-.61	.56	L	-.61	.56	-.21	.72	-.29	75	.7749	.0744	.7850	1.17	16	S16
P	-1.94	.60	L	1.29	.52	-.35	.80	-.44	75	.6603	.7500	.3865	.22	17	S17
P	-.61	.56	L	-1.95	.64	-1.34	.85	-1.59	75	.1165	3.5854	.0583	1.18	18	S18
P	1.98	.37	L	1.74	.37	-.23	.53	-.45	75	.6570	.0886	.9262	.22	19	S19
P	-.17	.47	L	-.81	.46	-.98	.66	-1.49	75	.3405	.9655	.3258	1.06	20	S20

³ William J Boone, John R S, and Melissa S Yale, *RASCH ANALYSIS IN THE HUMAN SCIENCES* (Springer, 2013).
⁴ Trevor G Bond and Christine M Fox, *APPLYING THE RASCH MODEL: FUNDAMENTAL MEASUREMENT IN THE HUMAN SCIENCES* (Psychology Press, 2013).
⁵ Boone, Staver, and Yale, *RASCH ANALYSIS IN THE HUMAN SCIENCES*.

Dalam hal ini file data mentah yang digunakan, dimasukkan dua data demografi yaitu jenis kelamin (Laki-laki, perempuan) serta asal sekolah (SD Ma'arif Sumberejo, SD Ma'arif NU Ngampir, dan SD Ma'arif Riyadul Arkham). Artinya terdapat tiga variabel demografi yang bisa digunakan untuk deteksi bias ini. Ketentuan jika diperoleh nilai probabilitas butir berada di bawah 5% (0,05), maka soal dikatakan mengandung bias. Lihat tabel 3 di atas butir soal yang mengandung bias terdapat pada butir soal S13, dimana butir soal S13 nilai probabilitas butirnya sebesar 0,02 yang mana berada di bawah 5% (0,05).

Cara mendeteksi adanya butir soal yang dikatakan bias (DIF), selain dengan angka, bisa juga ditampakan dengan adanya grafik yang dapat diperoleh secara langsung dari program Ministep dan hasilnya dalam bentuk file Microsoft Excel, seperti berikut ini:

Gambar 1. Grafik DIF



Tampak grafik yang mendekati batas atas pada pada grafik 1 di atas seperti halnya pada butir soal (S14), hal ini dideteksi adanya butir soal yang sulit, sedangkan pada grafik yang berada di bawah seperti halnya butir soal (S17), maka dari itu menunjukkan butir soal yang dikategorikan mudah. Informasi mendeteksi pada grafik di atas tentang butir soal DIF, untuk butir soal (S13) yang mempunyai (DIF) atau bias, tampak soal tersebut mudah dikerjakan oleh siswa Laki-laki, dibandingkan dikerjakan oleh siswa Perempuan. Untuk soal-soal lainnya berkaitan dengan perbedaan kemampuan siswa mengerjakan butir soal dengan benar tidak jauh berbeda.

KESIMPULAN

Hasil analisis butir soal untuk tes standarisasi mutu sekolah dasar mata pelajaran PAI dengan menggunakan pendekatan Rasch Model di SD Ma'arif NU Kecamatan Pandaan Kabupaten Pasuruan menunjukkan bahwa, pertama tingkat kesulitan butir soal (*item measure*) terlihat pada butir soal (S14), hal ini ditunjukkan dengan nilai Measure-nya adalah +2.82 dan hanya 25 siswa yang bisa menjawab dengan benar. Kedua tingkat kesesuaian butir soal (*item fit*) terdapat satu (1) respon yang berada di luar batas ruang kepercayaan Outfit yaitu pada butir soal (S14), butir soal selainya sudah memenuhi kriteria outfit z-standard, nilai outfit means-square, serta point measure correlation. Ketiga dideteksi

adanya butir soal yang bias yakni terdapat pada butir soal (S13), dimana butir soal tersebut nilai probabilitas butirnya sebesar 0,02 berada di bawah 5% (0,05).

UCAPAN TERIMAKASIH

Ucapan terima kasih disampaikan kepada DRPM Kemenristekdikti LL DIKTI Wilayah VII yang telah mendanai penelitian ini melalui Program Dosen Pemula tahun 2021 dan LPPM Universitas Yudharta Pasuruan yang telah membantu peneliti hingga selesai.

DAFTAR PUSTAKA

- Sumintono, Bambang., Widhiarso, Wahyu. *Aplikasi pemodelan Rasch: pada assessment pendidikan*. Trim komunikata, 2015.
- Bond, Trevor G, and Christine M Fox. *Applying the rasch model; fundamental measurement in the human sciences*. Psychology Press, 2013.
- Boone, William J, John R Staver, and Melissa S Yale. *Rasch analysis in the human sciences*. Springer, 2013.
- Isnani, Isnani, Purwo Susongko, Wikan Budi Utami, and Herani Tri Lestiani. "Estimation of College Students' Ability on Real Analysis Course Using Rasch Model." *Research and evaluation in education*, 5, No. 2 (2019): 95–102.
- Kadir, Abdul. "Menyusun Dan Menganalisis Tes Hasil Belajar." *Al-Ta'dib; Jurnal kajian ilmu kependidikan* 8, No. 2 (2015): 70–81.